

OIPE

RAW SEQUENCE LISTING

DATE: 05/14/2001

PATENT APPLICATION: US/09/847,670

TIME: 12:42:47

Input Set : A:\26862631.app

Output Set: N:\CRF3\05142001\I847670.raw

ENTERED

3 <110> APPLICANT: PHARMACIA & UPJOHN COMPANY
 5 <120> TITLE OF INVENTION: HEPATITIS C VIRUS HELICASE CRYSTALS, CRYSTALLOGRAPHIC
 6 STRUCTURE AND METHODS
 8 <130> FILE REFERENCE: 6263.N
 C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/847,670
 11 <141> CURRENT FILING DATE: 2001-05-02
 13 <150> PRIOR APPLICATION NUMBER: 60/201,598
 14 <151> PRIOR FILING DATE: 2000-05-03
 16 <160> NUMBER OF SEQ ID NOS: 3
 18 <170> SOFTWARE: PatentIn Ver. 2.1
 20 <210> SEQ ID NO: 1
 21 <211> LENGTH: 473
 22 <212> TYPE: PRT
 23 <213> ORGANISM: Hepatitis C virus
 25 <400> SEQUENCE: 1
 26 Met Val Asp Phe Ile Pro Val Glu Asn Leu Glu Thr Thr Met Arg Ser
 27 1 5 10 15
 29 Pro Val Phe Thr Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser Phe
 30 20 25 30
 32 Gln Val Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys
 33 35 40 45
 35 Val Pro Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn
 36 50 55 60
 38 Pro Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala
 39 65 70 75 80
 41 His Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr
 42 85 90 95
 44 Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly
 45 100 105 110
 47 Gly Cys Ser Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His
 48 115 120 125
 50 Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln
 51 130 135 140
 53 Ala Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala Thr Ala Thr Pro
 54 145 150 155 160
 56 Pro Gly Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu
 57 165 170 175
 59 Ser Thr Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu
 60 180 185 190
 62 Val Ile Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Lys Lys
 63 195 200 205
 65 Cys Asp Glu Leu Ala Ala Lys Leu Val Ala Leu Gly Ile Asn Ala Val
 66 210 215 220
 68 Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp
 69 225 230 235 240
 71 Val Val Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp
 72 245 250 255

RAW SEQUENCE LISTING

DATE: 05/14/2001

PATENT APPLICATION: US/09/847,670

TIME: 12:42:47

Input Set : A:\26862631.app

Output Set: N:\CRF3\05142001\I847670.raw

```

74 Phe Asp Ser Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val Asp
75           260           265           270
77 Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu Thr Ile Thr Leu Pro Gln
78           275           280           285
80 Asp Ala Val Ser Arg Thr Gln Arg Arg Gly Arg Thr Gly Arg Gly Lys
81           290           295           300
83 Pro Gly Ile Tyr Arg Phe Val Ala Pro Gly Glu Arg Pro Ser Gly Met
84 305           310           315           320
86 Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr Asp Ala Gly Cys Ala Trp
87           325           330           335
89 Tyr Glu Leu Thr Pro Ala Glu Thr Thr Val Arg Leu Arg Ala Tyr Met
90           340           345           350
92 Asn Thr Pro Gly Leu Pro Val Cys Gln Asp His Leu Glu Phe Trp Glu
93           355           360           365
95 Gly Val Phe Thr Gly Leu Thr His Ile Asp Ala His Phe Leu Ser Gln
96           370           375           380
98 Thr Lys Gln Ser Gly Glu Asn Leu Pro Tyr Leu Val Ala Tyr Gln Ala
99 385           390           395           400
101 Thr Val Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser Trp Asp Gln Met
102           405           410           415
104 Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu His Gly Pro Thr Pro
105           420           425           430
107 Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu Ile Thr Leu Thr His
108           435           440           445
110 Pro Val Thr Lys Tyr Ile Met Thr Cys Met Ser Ala Asp Leu Glu Val
111           450           455           460
113 Val Thr Ser His His His His His His
114 465           470
117 <210> SEQ ID NO: 2
118 <211> LENGTH: 9
119 <212> TYPE: PRT
120 <213> ORGANISM: Hepatitis C virus
122 <400> SEQUENCE: 2
123 Pro Thr Gly Ser Gly Lys Ser Thr Lys
124 1           5
127 <210> SEQ ID NO: 3
128 <211> LENGTH: 8
129 <212> TYPE: PRT
130 <213> ORGANISM: Artificial Sequence
132 <220> FEATURE:
133 <223> OTHER INFORMATION: Description of Artificial Sequence: conserved
134     sequence motif VI wherein Xaa represents
135     unconserved amino acids within the motif
137 <400> SEQUENCE: 3
W--> 138 Gln Arg Xaa Gly Arg Xaa Gly Arg
139 1           5

```

VERIFICATION SUMMARY

DATE: 05/14/2001

PATENT APPLICATION: US/09/847,670

TIME: 12:42:48

Input Set : A:\26862631.app

Output Set: N:\CRF3\05142001\I847670.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application Number

L:138 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:3

L:138 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:3

L:138 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3